

Sequential Joint Spectrum Sensing and Channel Estimation for Dynamic Spectrum Access

Yasin Yilmaz*, Ziyu Guo[†], and Xiaodong Wang*

Abstract

Dynamic spectrum access under channel uncertainties is considered. With the goal of maximizing the secondary user (SU) throughput subject to constraints on the primary user (PU) outage probability we formulate a joint problem of spectrum sensing and channel state estimation. The problem is cast into a sequential framework since sensing time minimization is crucial for throughput maximization. In the optimum solution, the sensing decision rule is coupled with the channel estimator, making the separate treatment of the sensing and channel estimation strictly suboptimal. Using such a joint structure for spectrum sensing and channel estimation we propose a distributed (cooperative) dynamic spectrum access scheme under statistical channel state information (CSI). In the proposed scheme, the SUs report their sufficient statistics to a fusion center (FC) via level-triggered sampling, a nonuniform sampling technique that is known to be bandwidth-and-energy efficient. Then, the FC makes a sequential spectrum sensing decision using local statistics and channel estimates, and selects the SU with the best transmission opportunity. The selected SU, using the sensing decision and its channel estimates, computes the transmit power and starts data transmission. Simulation results demonstrate that the proposed scheme significantly outperforms its conventional counterparts, under the same PU outage constraints, in terms of the achievable SU throughput.

Index Terms: sensing-based dynamic spectrum access, sequential joint detection and estimation, cooperative dynamic spectrum access, level-triggered sampling

I. INTRODUCTION

Addressing the well-known problem of spectrum utilization scarcity in current wireless networks, the cognitive radio (CR) technology employs a hierarchical spectrum access model consisting of primary users (PUs) and secondary users (SUs) [1]. In this model, both PUs and SUs are able to access a same

*Electrical Engineering Department, Columbia University, New York, NY 10027.

[†]National Mobile Communications Research Lab., Southeast University, Nanjing 210096, China.

band with a higher priority for PUs. The spectrum sharing between PUs and SUs can be realized in an *underlay* fashion, which allows SUs to coexist with PUs without sensing the spectrum band. Thus, SUs are blind to the idle state of PUs (spectrum holes), resulting in a worst-case assumption that PUs use the band all the time. As a result, SUs can coexist only with severe constraints on the transmission power in order to protect the quality of service (QoS) of PUs. Focusing on the analysis of underlay spectrum access, [2]–[4] derive fading channel capacities and optimum power allocation strategies for SUs. In contrast to underlay, the *opportunistic access* approach permits the existence of SUs only when PUs are idle, i.e., no coexistence. Hence, in this approach there is no harsh constraints on the SU transmission power. Instead, an effective spectrum sensing scheme is needed [5]–[7]. In [5], [6] the SU throughput is maximized while satisfying the PU QoS constraints.

Methods for combining the underlay and opportunistic access approaches have also been proposed, e.g., [8]–[11]. In such combined methods, the SU senses the spectrum band, as in opportunistic access, and controls its transmit power using the sensing result, which allows SU to coexist with PU, as in underlay. While deriving the power control function, the average or peak constraints on SU transmit power and PU interference level are imposed [10], [11]. In this paper, we propose such a combined method under the peak interference and power constraints. In spectrum access methods it is customary to assume perfect channel state information (CSI) at the SU, e.g., [2]–[4], [8]–[11]. That is, the perfect CSI of SU channels (and even PU channels) can be made available to the SU. The quantized CSI case is treated in [11]. However, how to obtain the CSI in the process of dynamic spectrum access has not been addressed. We consider the problem of joint spectrum sensing and channel estimation in this work.

For such a joint problem, a straightforward solution is to treat the two subproblems separately by using the optimum solution for each subproblem. More specifically, one can use the likelihood ratio test (LRT) for spectrum sensing and the minimum mean square error (MMSE) estimator for channel estimation to solve the joint problem. However, as shown in [12], [13], treating each subproblem separately and solving it optimally does not necessarily result in the optimum overall performance. In [12], [14], [15], optimum solutions to different formulations of the joint detection and estimation problem are given under the fixed-sample-size framework. More recently, in [13] a *sequential joint detection and estimation* problem is considered, and the optimum solution is given, where the decision rule is a function of the estimator, making the separate treatment strictly suboptimal. The sequential framework ideally suits the goal of maximizing the SU throughput in dynamic spectrum access. In particular, it is desirable to perform reliable sensing as soon as possible to let the SU transmit data as long as possible, leading to higher throughput. Indeed, in the sequential framework the sensing time is minimized. Here we propose

a dynamic spectrum access method based on sequential joint spectrum sensing and channel estimation.

Pilot signals are often used in channel estimation, e.g., [16], [17], and also in spectrum sensing, e.g., [18]–[20]. We similarly propose to make use of the pilot signals transmitted for PU communications to jointly sense and estimate the channels linked to the SU. In a cognitive radio network, multiple SUs can cooperate to sense the spectrum by sharing their local information either over a fusion center (FC) or directly with other SUs. For such a decentralized system bandwidth and energy-efficient scheme is required for information transmission and processing. Recently, in a series of papers [21]–[23], it is shown that a nonuniform sampling technique called *level-triggered sampling* is an ideal fit for distributed information transmission and processing. This is because it enables highly accurate recovery at the FC by transmitting only a single bit per sample. Furthermore, it allows for complete asynchrony among SUs, a highly desirable feature in distributed systems, and censors uninformative local information. Due to its attractive features we use level-triggered sampling in the proposed dynamic spectrum access scheme to enable cooperation between SUs.

The remainder of the paper is organized as follows. In Section II, we formulate the problem and briefly discuss the conventional spectrum access methods. Then, in Section III the sequential joint spectrum sensing and channel estimation problem is introduced and the optimum solution is given. The proposed cooperative spectrum access scheme is given in Section IV, and simulation results comparing its performance with other schemes are provided in Section V. Finally, the paper is concluded in Section VI.

II. SYSTEM DESCRIPTIONS

Consider a cognitive radio network consisting of a primary user (PU) pair, a secondary user transmitter (SU Tx) and receiver (SU Rx), and a fusion center (FC), as shown in Fig. 1, where the PU pair can simultaneously communicate to each other through full duplexing. Although no direct communication takes place between the PUs and the SUs, interference to the PU communications occurs through the cross links, represented by dashed lines in Fig. 1. The FC facilitates cooperation among the SUs, and it can be either a dedicated entity or one of the SUs. The channel, i.e., cross link, between PU i and SU Tx is represented by a complex random coefficient, i.e., channel gain, h_{i1} . Similarly the complex random coefficient h_{i2} denotes the channel gain between the PU i and SU Rx. We assume Rician fading channels, i.e., the real and imaginary parts of h_{ik} , $\Re(h_{ik})$ and $\Im(h_{ik})$, are independent and identically distributed (i.i.d.) as $\mathcal{N}(\mu_{ik}, \sigma_{ik}^2)$, $\forall i, k$, with $\mu_{ik} = 0$ corresponding to Rayleigh fading channels. Moreover, $\{h_{ik}\}$ are assumed to be independent, but they are in general not identically distributed with different means

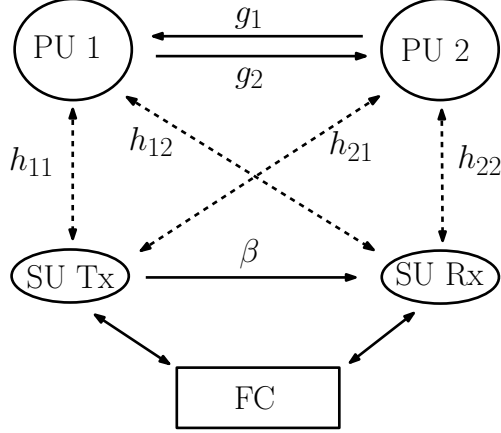


Fig. 1. The cognitive radio system under consideration.

and variances.

A. Problem Formulation

As a fundamental requirement in cognitive radio systems, the SUs should not cause degradation in the quality of service (QoS) to the PUs. In other words, the interference from the SUs to the PUs must be kept below some maximum tolerable levels. Under such interference constraints, a natural objective is to maximize the SU throughput, i.e., the average bit-rate of SU Tx. Hence, assuming Gaussian noise in channels between the PUs and also between the SUs we aim to solve the following optimization problem

$$\max_{P(h_{11}, h_{21}) \leq P_{\max}} \begin{cases} \log \left(1 + \frac{|\beta|^2 P(h_{11}, h_{21})}{N_0} \right) & \text{if } H_0 \\ \log \left(1 + \frac{|\beta|^2 P(h_{11}, h_{21})}{N_0 + |h_{12}|^2 Q_1 + |h_{22}|^2 Q_2} \right) & \text{if } H_1 \end{cases}, \quad (1)$$

$$\text{s.t. } |h_{11}|^2 P(h_{11}, h_{21}) \leq I_1 \text{ and } |h_{21}|^2 P(h_{11}, h_{21}) \leq I_2$$

where $P(h_{11}, h_{21})$ is the transmit power of SU Tx, constrained by the maximum power P_{\max} , and is a function of the channel gains h_{11}, h_{21} , between SU Tx and the PUs; β and N_0 are the channel gain and the variance of the Gaussian noise, respectively, between SU Tx and SU Rx; Q_1 and Q_2 are the transmit powers for PUs; and I_1 and I_2 are the maximum tolerable interference powers at PUs, which are determined by the PU outage constraints. The null hypothesis H_0 and the alternative hypothesis H_1 correspond to the absence and presence of PU communication, respectively. More specifically, $Q_1 = Q_2 = 0$ under H_0 , whereas $\max\{Q_1, Q_2\} \neq 0$ under H_1 .

In (1), we in fact maximize the average capacity of a Gaussian channel, where the interference constraint

I_1 is determined according to the outage constraint on another Gaussian channel g_1

$$\mathbb{P}\left(\log\left(1 + \frac{|g_1|^2 Q_2}{\eta_1 + \underbrace{|h_{11}|^2 P}_{I_1}}\right) < R_2\right) \leq P_{\text{out}}, \quad (2)$$

where η_1 is the variance of the Gaussian noise; and R_2 is the bit-rate of PU 2. The outage constraint in (2) yields the interference constraint in (1), given $Q_2, R_2, \eta_1, P_{\text{out}}$. The maximum interference value I_2 is written similarly. We assume I_1 and I_2 are available to SUs. In a careful design, there should be some safety margin between the probability on the left hand-side of (2) and P_{out} while determining I_i . This is because SUs may unintentionally exceed I_i due to lack of information on the true hypothesis and the actual channel coefficients.

B. Spectrum Access Methods

The conventional spectrum access methods for cognitive radio, namely the opportunistic access and underlay methods, provide simplistic solutions to (1). In particular, the opportunistic access method focuses only on the binary hypothesis test, i.e., spectrum sensing, and conforms to the interference constraints by simply turning off SU Tx, i.e., $P = 0$, when H_1 is declared. When H_0 is declared, SU Tx transmits at the maximum power, i.e., $P = P_{\text{max}}$. On the other hand, the underlay method does not perform spectrum sensing and solves only the constrained optimization problem under H_1 . As a result, the constant power $P = \min\left\{P_{\text{max}}, \frac{I_1}{|h_{11}|^2}, \frac{I_2}{|h_{21}|^2}\right\}$ is transmitted under both H_0 and H_1 . It is seen that deep fades in the cross links $\{h_{ik}\}$ are beneficial for the SU throughput.

In practice, the channels g_1, g_2 between the PUs, and the cross links $\{h_{ik}\}$ are not known a priori. Hence, the PUs perform a preamble communication with duration T_p at the beginning of each data transmission frame to estimate g_1 and g_2 . Specifically, for transmission frame m , as shown in Fig. 2, PU i estimates g_i during $t \in (T(m-1), T(m-1) + T_p]$ using pilot symbols, and then data transmission takes place during $t \in (T(m-1) + T_p, Tm]$, where T is the frame duration. Assuming the SUs are synchronized with the PU frame timing and observe pilot signals, each SU can estimate its cross links during each preamble period.

As opposed to the naive solutions of the conventional spectrum access methods, an efficient solution to (1) should involve both spectrum sensing and channel estimation, hence it is a combination of the opportunistic access and underlay methods. For example, at a fixed time $\tau \in (0, T_p]$, we can employ the optimum detector, i.e., the likelihood ratio test (LRT) for spectrum sensing, and the optimum estimator, i.e., the minimum mean square error (MMSE) estimator for channel estimation. Once we obtain the

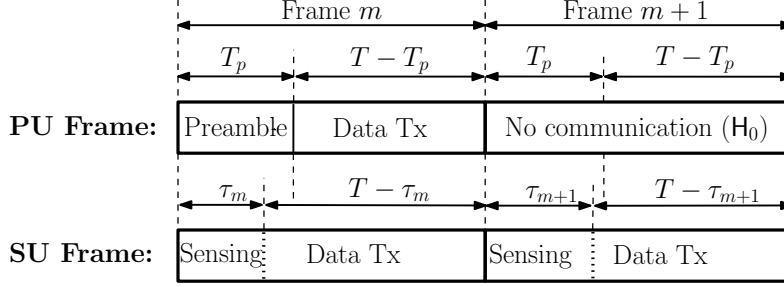


Fig. 2. Frame structures for the PU communication and the SU communication.

spectrum sensing result (H_0 or H_1) and the channel estimates $\{\hat{h}_{ik}\}$, we can use them to solve (1) as follows: SU Tx transmits with $P = P_{\max}$ when H_0 is declared, as in opportunistic access, and with $P = \min \left\{ P_{\max}, \frac{I_1}{|\hat{h}_{11}|^2}, \frac{I_2}{|\hat{h}_{21}|^2} \right\}$ when H_1 is declared, as in underlay.

As a more sophisticated example, instead of performing fixed-sample-size detection and estimation (at a fixed time τ) we can determine the sample number based on the observed samples, resulting in a *sequential* method with a random sensing time τ . In particular, we can use the sequential probability ratio test (SPRT) [24], which is the optimum sequential detector for i.i.d. observations in terms of minimizing the average detection delay, for spectrum sensing, and then use the MMSE estimator at the random sensing time τ to estimate the unknown channel gains $\{h_{ik}\}$.

However, the above approaches based on separate detection and estimation in general may not yield the optimal solution. In the following section, we propose a new and powerful solution based on sequential joint detection and estimation.

III. SEQUENTIAL JOINT SPECTRUM SENSING AND CHANNEL ESTIMATION

In this section, we focus on SU Tx, to introduce the sequential joint spectrum sensing and channel estimation framework. Hence, the subscript denoting SU Tx is dropped.

A. Motivation

In our system model, in each frame m during the preamble period $t \in (T(m-1), T(m-1) + T_p]$, the signal received by the SU from PU i is given by

$$y_i[t] = \begin{cases} w_i[t] & \text{if } H_0 \\ h_i p_i[t] + w_i[t] & \text{if } H_1 \end{cases}, \quad i = 1, 2, \quad t = 1, 2, \dots, \quad (3)$$

where $w_i[t] \sim \mathcal{N}_c(0, N_0^i)$ is the complex additive white Gaussian noise; $h_i \sim \mathcal{N}_c(\mu_i, \sigma_i^2)$ is the proper complex channel coefficient between PU i and the SU; and $p_i[t]$ is the complex random pilot signal

used in the preamble. The processes $\{p_i[t]\}$ and $\{w_i[t]\}$ are independent and they are independent of the random variable h_i . We assume the SU observes $p_i[t]$ at time t , e.g., the SU knows the seed of the random number generator that generates $p_i[t]$, for $i = 1, 2$.

In (3), we would like to decide between H_0 and H_1 as soon as possible, and also estimate h_i if we decide for H_1 . In other words, our objective is to have a reliable estimate of the channel coefficient h_i every time we detect the presence of PU communication. Deciding as soon as possible is important because an early sensing time, i.e., small τ , enables the SU to transmit data for a longer period of time, i.e., large $T - \tau$, increasing the SU throughput. On the other hand, the SU transmit power, which is a function of sensing decision and estimates of $\{h_1, h_2\}$, should obey the PU maximum interference constraints. Small τ may increase the misdetection probability and decrease the estimation accuracy, leading to the violation of such constraints and PU outage. Hence, there is a tradeoff in selecting the τ value. Conventionally τ is selected offline, resulting in a fixed-sample-size test. Whereas in a sequential test τ is determined online, i.e., it depends on the observations, and thus it is random. Although sequential tests are more sophisticated than fixed-sample-size tests, they are much more powerful in minimizing the average sensing time, $E[\tau]$, hence suit better the cognitive radio application.

In the separate detection and estimation approach, the unknown channel gain h_i is treated as a nuisance parameter while performing detection. However, channel estimation is an integral part of the problem of interest. Hence, formulating the problem as a joint detection and estimation problem is a more natural way to obtain better overall performance, i.e., SU throughput. Indeed it was shown in [13] that the combined optimum detector and optimum estimator do not produce the optimum overall detection and estimation performance.

B. Problem Formulation

Since the results in [13] are obtained for real signals, for analytical convenience in our problem we will treat a complex observation (channel) as two real observations (channels). Specifically, we compute $y_i^1[t] \triangleq \Re(p_i[t]^* y_i[t])$ and $y_i^2[t] \triangleq \Im(p_i[t]^* y_i[t])$, hence instead of (3) we use the following signal model

$$y_i^n[t] = \begin{cases} w_i^n[t] & \text{if } H_0 \\ h_i^n |p_i[t]|^2 + w_i^n[t] & \text{if } H_1 \end{cases}, \quad i = 1, 2, \quad n = 1, 2, \quad t = 1, 2, \dots, \quad (4)$$

where $h_i^1 \triangleq \Re(h_i)$, $h_i^2 \triangleq \Im(h_i)$, $w_i^1[t] \triangleq \Re(p_i[t]^* w_i[t])$, and $w_i^2[t] \triangleq \Im(p_i[t]^* w_i[t])$. Note in (4) that $h_i^n \sim \mathcal{N}\left(\frac{\mu_i}{2}, \frac{\sigma_i^2}{2}\right)$, $n = 1, 2$; and given $p_i[t]$, the noise $w_i^n[t] \sim \mathcal{N}\left(0, |p_i[t]|^2 \frac{N_0}{2}\right)$, $n = 1, 2$, and $\{w_i^n[t]\}$ are independent across channels (for different i) and time. Similar to (3), we want to sequentially decide between H_0 and H_1 , and also estimate h_i^n when we decide on H_1 . To present the sequential joint detection

and estimation (SJDE) problem and the optimum solution to it we first focus on a single channel case, i.e., the signal model in (4) for specific i, n values. In particular, the SU, using its observations $\{(y_i^n[t], p_i[t])\}_t$ through the real channel n linked to PU i , wants to jointly detect the PU communication and estimate the channel coefficient h_i^n when it decides on its presence.

In sequential methods, in general, the average sample number, which corresponds to the average sensing time in our context, is minimized subject to a set of constraints, e.g., false alarm and misdetection constraints for detection, and mean squared error constraint for estimation. In the proposed joint framework we use the following combined cost function

$$\mathcal{C}(\tau, d_\tau, \hat{x}_\tau) = c_0 P_0(d_\tau = 1 | \mathcal{F}_\tau) + c_1 P_1(d_\tau = 0 | \mathcal{F}_\tau) + c_e \bar{E}_1 \left[(\hat{x}_\tau - x)^2 \mathbb{1}_{\{d_\tau=1\}} + x^2 \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right] \quad (5)$$

where d_τ is the decision function; $x \triangleq h_i^n$ is the unknown parameter; \hat{x}_τ is the estimate of x ; c_0, c_1, c_e are nonnegative constants selected by the designer; P_0 and E_0 denote the probability measure and expectation under hypothesis H_0 ; \bar{P}_1 and \bar{E}_1 denote the probability measure and expectation under H_1 ; P_1 and E_1 denote the probability measure and expectation under H_1 with x being marginalized; $\mathcal{F}_t = \sigma\{p_i[1], \dots, p_i[t]\}$ is the σ -algebra, that is, the accumulated history pertinent to the observed process $\{p_i[t]\}$; and $\mathbb{1}_{\{A\}}$ is the indicator of the event A , taking the value 1 if A occurs and 0 otherwise. Then, our constrained optimization problem is given by

$$\min_{\tau, d_\tau, \hat{x}_\tau} E[\tau | \mathcal{F}_\tau] \quad \text{subject to} \quad \mathcal{C}(\tau, d_\tau, \hat{x}_\tau) \leq \alpha, \quad (6)$$

where $\alpha > 0$ is a given constant, denoting the target accuracy level.

Our formulation in (5) and (6) is conditioned on the auxiliary statistic \mathcal{F}_t because using such extra information we can assess the accuracy of the detector and estimator more precisely than the unconditional formulation. More specifically, since $P_0(d_\tau = 1) = E_0[\mathbb{1}_{\{d_\tau=1\}}] = E[E_0[\mathbb{1}_{\{d_\tau=1\}} | \mathcal{F}_\tau]] = E[P_0(d_\tau = 1 | \mathcal{F}_\tau)]$, there is no need to use the expectation, e.g., $P_0(d_\tau = 1)$, of an accuracy assessment term when the term itself, e.g., $P_0(d_\tau = 1 | \mathcal{F}_\tau)$, is available. Moreover, with the conditional formulation used in (5) we do not need to specify the distribution of the pilot signal $p_i[t]$. Note that the constraint $\mathcal{C}(\tau, d_\tau, \hat{x}_\tau) \leq \alpha$ in (6) is required to hold for each realization of the process $\{p_i[t]\}$, hence is stricter than its unconditional counterpart, which is required to hold only on average with respect to $\{p_i[t]\}$.

In (5), the first two terms, which are related to the detection problem, correspond to the false alarm and misdetection probabilities (P_f and P_m), respectively. On the other hand, the last term, which is related to the estimation problem, depends on both the decision and estimation strategies. Without this term, i.e., for $c_e = 0$, the combined cost depends only on the decision function d_τ , implying that the joint problem

reduces into a pure detection problem.

Similar to \mathcal{F}_t let $\mathcal{G}_t = \sigma\{(y_i^n[1], p_i[1]), \dots, (y_i^n[t], p_i[t])\}$ denote the σ -algebra generated by the processes $\{y_i^n[t]\}$ and $\{p_i[t]\}$, i.e., the complete observation history. Then, we have the corresponding filtrations $\{\mathcal{F}_t\}_{t \geq 0}$ and $\{\mathcal{G}_t\}_{t \geq 0}$. In general, the solution we seek should use all available information, that is, we are looking for a triplet $(\tau, d_\tau, \hat{x}_\tau)$ where τ is $\{\mathcal{G}_t\}$ -adapted, d_τ and \hat{x}_τ are \mathcal{G}_τ -measurable. It is known in the pure estimation problem that with a $\{\mathcal{G}_t\}$ -adapted stopping time τ , in most cases, finding an optimum sequential estimator (τ, \hat{x}_τ) is not tractable [25]. Instead, [26] considered using an $\{\mathcal{F}_t\}$ -adapted stopping time, which was later shown to have a simple optimal solution for continuous-time and discrete-time observations in [27] and [23], respectively. Similarly, in the pure detection problem with a $\{\mathcal{G}_t\}$ -adapted τ we have a two-dimensional optimal stopping problem, which is not tractable. Consequently, following the approach used for the pure estimation problem in [23], [26], [27] we consider $\{\mathcal{F}_t\}$ -adapted stopping times for our joint problem. On the other hand, we are still interested in \mathcal{G}_τ -measurable decision rule d_τ and estimator \hat{x}_τ , which use all available information acquired up to stopping time τ . As a result, the problem in (6) takes the following form

$$\min_{\tau, d_\tau, \hat{x}_\tau} \tau \quad \text{s.t.} \quad \mathcal{C}(\tau, d_\tau, \hat{x}_\tau) \leq \alpha. \quad (7)$$

C. The Optimal Solution

The following theorem gives the optimum solution to the above problem.

Theorem 1. *Consider the observations $\{(y_i^n[t], p_i[t])\}_t$ obtained through the real channel $x = h_i^n$. Then, the optimum triplet $(\tau, d_\tau, \hat{x}_\tau)$ of stopping time, decision function, and estimator for the sequential joint detection and estimation (SJDE) problem in (7) is given by*

$$\tau = \min \{t > 0 : U_t^i \geq \gamma\} \quad (8)$$

$$d_\tau = \begin{cases} 1 & \text{if } L_\tau^{in} \geq \log \frac{c_0}{c_1 + c_e \hat{x}_\tau^2} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$\hat{x}_t = \frac{V_t^{in} + \frac{\mu_i}{2} \frac{N_0^i}{\sigma_i^2}}{U_t^i + \frac{N_0^i}{\sigma_i^2}}, \quad (10)$$

where $U_t^i \triangleq \sum_{m=1}^t |p_i[m]|^2$ is the conditional Fisher information given \mathcal{F}_t in estimating $x = h_i^n$ under H_1 [cf. (4)]; $V_t^{in} \triangleq \sum_{m=1}^t y_i^n[m]$; γ is a constant threshold [13, Theorem 1]; and

$$L_t^{in} \triangleq \frac{\left(V_t^{in} + \frac{\mu_i}{2} \frac{N_0^i}{\sigma_i^2}\right)^2}{N_0^i \left(U_t^i + \frac{N_0^i}{\sigma_i^2}\right)} - \frac{\mu_i^2}{4\sigma_i^2} - \frac{1}{2} \log \left(\frac{\sigma_i^2}{N_0^i} U_t^i + 1 \right) \quad (11)$$

is the conditional log-likelihood ratio (LLR) between the hypotheses H_0 and H_1 given \mathcal{F}_t with x under H_1 being marginalized [13, Lemma 2].

Proof: The proof closely follows [13], so omitting the details we only highlight the differences here. The main difference is that the noise in (4) is independent across time but has a time-varying variance, whereas i.i.d. noise is assumed in [13]. The common term $|p_i[t]|^2$ in the variance and the mean of the observation $y_i^n[t]$ given h_i^n and $p_i[t]$ under H_1 cancels while writing the estimator \hat{x}_t and the LLR L_t^{in} . As a result, the definitions of the Fisher information term U_t^i and its companion V_t^{in} differ from their counterparts in [13]. However, the results in [13] still hold here with the new definitions of U_t^i, V_t^{in} and the noise variance appearing without $|p_i[t]|^2$ as $\frac{N_0}{2}$ since U_t^i, V_t^{in} and accordingly other key terms maintain their properties, e.g., U_t^i is increasing. ■

The optimum stopping rule in (8) terminates getting new samples when the conditional Fisher information exceeds a threshold whose exact expression can be found in [13, Theorem 1]. Since the conditional Fisher information is increasing, it is guaranteed to have a finite stopping, i.e., sensing, time. The optimum decision function in (9) is a modification of the well-known likelihood ratio test (LRT). For $c_e = 0$, i.e., in the pure detection problem, it boils down to LRT. For $c_e \neq 0$ the estimator is incorporated into LRT. The way it modifies LRT is quite intuitive. When the estimate is nonzero, the threshold is decreased, supporting a decision in favor of H_1 . The further the estimate is from zero, the easier to decide for H_1 . The estimate provides some side information about the true hypothesis, and the optimum solution to the joint problem uses it. Such a plausible modification appears in the decision function since the detection and estimation problems are formulated jointly. The optimum estimator, given in (10), is the minimum mean square error (MMSE) estimator, which is equivalent to the maximum a posteriori (MAP) estimator in the Gaussian case under consideration.

D. Discussions

Comparing the optimum triplet in Theorem 1 with the combined SPRT & MMSE method, we see that there are fundamental differences in the stopping rule and decision function. In SPRT [28], the stopping time and detection decision are determined together through a common procedure. More specifically, two thresholds are used to jointly terminate the scheme and make a decision. When the scheme terminates, the decision is already clear as it is determined by the threshold that causes termination. As a result, the performance metrics $P_f, P_m, E[\tau]$, and also $MSE = E[(\hat{x}_\tau - x)^2]$ are closely interrelated since they are all controlled by the two thresholds, which are the only system parameters. On the other hand, in SJDE the

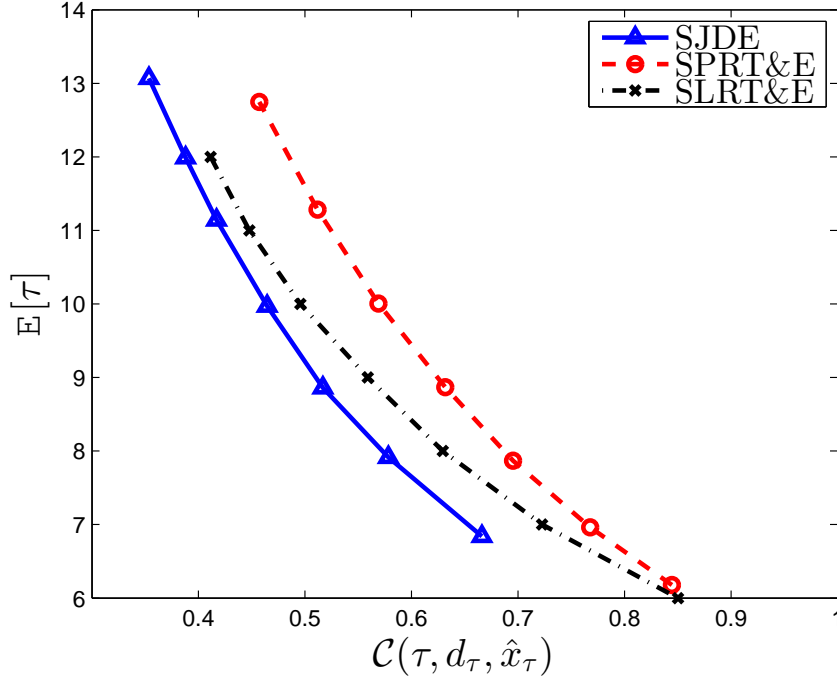


Fig. 3. Average sensing time vs. combined cost for the SJDE in Theorem 1, and the SPRT & MMSE and the sequential LRT & MMSE equipped with the stopping rule of SJDE.

stopping time and decision are computed using two separate procedures. First the stopping time is found by performing a single-threshold-test, and then the decision is made via a modified LRT. In particular, $E[\tau]$ and $E[(\hat{x}_\tau - x)^2]$ are controlled by only the stopping threshold γ , whereas P_f and P_m are controlled by γ , c_0 , c_1 , and c_e . That is to say, $E[\tau]$ and $E[(\hat{x}_\tau - x)^2]$ can be controlled independently from P_f and P_m through γ , and similarly P_f and P_m can be controlled independently from $E[\tau]$ and $E[(\hat{x}_\tau - x)^2]$ through c_0 , c_1 , and c_e . The latter set of parameters enables a trade-off between P_f and P_m without affecting $E[\tau]$ and $E[(\hat{x}_\tau - h_i^n)^2]$. For instance, we can trade false alarm probability P_f for misdetection probability P_m , which is crucial for complying with the outage constraints of PUs, by decreasing the ratio of c_0 to c_1 or c_e without sacrificing early stopping or estimation quality. We obviously have a higher degree of freedom in SJDE than SPRT due to the number of parameters that control the system performance, which endows us with the ability to strike a right balance between our objectives of early stopping, and accurate detection and estimation.

In Fig. 3, we numerically show the superior performance of SJDE over the combined SPRT & MMSE (SPRT&E) in terms of the combined detection and estimation cost in (5). We also compare SJDE with the sequential LRT & MMSE (SLRT&E) that is equipped with the stopping rule of SJDE to demonstrate

the advantage of incorporating the estimate into the decision function. SLRT&E uses the unmodified (original) LRT to detect, hence can be seen as a separate-formulation-method. It outperforms SPRT&E since it enjoys the flexibility of SJDE to strike a desired balance for the specific problem of interest by employing two separate procedures, namely the stopping rule of SJDE and LRT, to terminate the scheme and make a decision respectively. In our problem of interest, it is crucial that SUs do not violate the maximum interference constraint, which in turn ensures an admissible PU outage probability. In case of misdetection the SU transmits with maximum power, which may cause the violation of outage constraint. Even when the SU correctly detects PU communication, poor channel estimate may still cause the SU to transmit with a non-admissible power. On the other hand, the false alarm, which corresponds to deciding on H_1 under H_0 , is not related to the outage constraint, but only degrades the SU throughput. Therefore, in the combined cost expression in (5) the second and third terms are more important than the first term. Accordingly, in Fig. 3 we use $c_0 = c_1 = 0.2$ and $c_e = 0.6$. Since the second part of the third term in (5) already penalizes misdetection, we do not differentiate between the coefficients, c_0 and c_1 , of the detection error probabilities. In Fig. 3, referring to (4) we use $\mu_i = 0$, i.e., Rayleigh fading channel h_i^n , and $\sigma_i^2 = N_0^i = \mathbb{E}[|p_i[t]|^2] = 1$.

E. SJDE for a Single SU with Multiple Channels

Here, following the optimum SJDE scheme in Theorem 1 for the single channel case we are interested in finding the optimum SJDE scheme for the SU observing the signals $\{y_i^n[t]\}$ and $\{p_i[t]\}$ through the channels $\{h_i^n\}$, $i = 1, 2$, $n = 1, 2$ from PU 1 and PU 2. We first need to modify the cost function in (5) by adding the new MSE terms, i.e.,

$$\begin{aligned} \mathcal{C}(\tau, d_\tau, \hat{h}_i^n[\tau]) &= c_0 P_0(d_\tau = 1 | \mathcal{F}_\tau) + c_1 P_1(d_\tau = 0 | \mathcal{F}_\tau) \\ &\quad + c_e \sum_{i=1}^2 \sum_{n=1}^2 \bar{\mathbb{E}}_1 \left[\left(\hat{h}_i^n[\tau] - h_i^n \right)^2 \mathbb{1}_{\{d_\tau=1\}} + (h_i^n)^2 \mathbb{1}_{\{d_\tau=0\}} \middle| \mathcal{F}_\tau \right]. \end{aligned} \quad (12)$$

The following theorem, whose proof is provided in the Appendix, gives the optimum SJDE scheme in this case.

Theorem 2. *With the cost function in (12), and the observations $\{y_i^n[t], p_i[t]\}$ obtained through the channels $\{h_i^n\}$, $i = 1, 2$, $n = 1, 2$ from PU 1 and PU 2, the optimum triplet of stopping time, decision function, and estimator for the sequential joint detection and estimation (SJDE) problem in (7) is given*

by

$$\tau = \min \{t > 0 : U_t \geq \bar{\gamma}\} \quad (13)$$

$$d_\tau = \begin{cases} 1 & \text{if } L_\tau \geq \log \frac{c_0}{c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 (\hat{h}_i^n[\tau])^2} \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$\hat{h}_i^n[t] = \frac{V_t^{in} + \frac{\mu_i}{2} \frac{N_0^i}{\sigma_i^2}}{U_t^i + \frac{N_0^i}{\sigma_i^2}}, \quad \forall i, n, \quad (15)$$

where $U_t = \sum_{i=1}^2 \sum_{m=1}^t |p_i[m]|^2$ is the conditional Fisher information given \mathcal{F}_t under H_1 ; $\bar{\gamma}$ is a constant threshold [cf. (8)]; and $L_t = \sum_{i=1}^2 \sum_{n=1}^2 L_t^{in}$ [cf. (11)] is the global LLR.

For systems with multiple SU pairs, in the next section we propose a distributed and cooperative spectrum access method which selects the SU with the maximum achievable throughput, and controls its transmit power.

IV. DISTRIBUTED SPECTRUM ACCESS BASED ON SJDE

In the previous section we formulated the joint spectrum sensing and channel estimation problem for a single SU and gave the optimal solution to it. In this section we consider K SUs, i.e., $K/2$ SU transmitter-receiver pairs, where each SU observes signals through 4 different real channels (2 from each PU). All observations of K SUs through $4K$ channels are used to detect a single event, namely the PU communication. Hence, under the joint framework introduced in Section III, SUs can cooperate to detect the PU communication. We next propose a bandwidth and energy-efficient distributed spectrum access algorithm for the cognitive radio system under consideration.

A. SJDE-based Spectrum Access with Multiple SUs

We now consider the multi-SU case for SJDE, and propose a dynamic spectrum access method (DSA-SJDE). From (12), we have the following cost function,

$$\begin{aligned} \mathcal{C}(\tau, d_\tau, \hat{h}_i^n[\tau]) &= c_0 P_0(d_\tau = 1 | \mathcal{F}_\tau) + c_1 P_1(d_\tau = 0 | \mathcal{F}_\tau) \\ &\quad + c_e \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 \bar{E}_1 \left[\left(\hat{h}_{ik}^n[\tau] - h_{ik}^n \right)^2 \mathbb{1}_{\{d_\tau=1\}} + (h_{ik}^n)^2 \mathbb{1}_{\{d_\tau=0\}} \middle| \mathcal{F}_\tau \right]. \end{aligned} \quad (16)$$

Note that all K SUs observe the same pilot signals $\{p_1[t]\}$ and $\{p_2[t]\}$. Hence, from (13) it is seen that they have the same stopping time, which in this case serves as a global stopping time. Each channel coefficient h_{ik}^n is again estimated using (15) for all k, i, n because they are independent. Since the observations

$\{y_{ik}^n[t]\}_k$, across SUs, are independent given $\{p_i[t]\}$, the global LLR is written as $L_t = \sum_{k=1}^K L_t^k$, and as in (14) we sum the channel estimates to write the threshold. Then, substituting the global LLR and the global threshold in (14) we obtain the decision function for the multi-SU case.

Corollary 1. *In the multi-SU case with the cost function in (16), the optimum solution to (7) is given by*

$$\tau = \min \{t > 0 : U_t \geq \bar{\gamma}\} \quad (17)$$

$$d_\tau = \begin{cases} 1 & \text{if } L_\tau \geq \log \frac{c_0}{c_1 + c_e \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 (\hat{h}_{ik}^n[\tau])^2} \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

$$\hat{h}_{ik}^n[t] = \frac{V_t^{ikn} + \frac{\mu_{ik}}{2} \frac{N_0^{ik}}{\sigma_{ik}^2}}{U_t^i + \frac{N_0^{ik}}{\sigma_{ik}^2}}, \quad \forall i, k, n, \quad (19)$$

where $U_t = \sum_{i=1}^2 \sum_{m=1}^t |p_i[m]|^2$ is the conditional Fisher information given \mathcal{F}_t under H_1 ; $\bar{\gamma}$ is a constant threshold [cf. (8)]; $V_t^{ikn} = \sum_{m=1}^t y_{ik}^n[m]$ [cf. (4)]; and $L_t = \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 L_t^{ikn}$ [cf. (11)] is the global LLR.

It looks like the SJDE scheme for the multi-SU case simply follows from (13)–(15) in the single-SU case. However, in the multi-SU case the stopping time, detector, and estimator are computed at the FC, which requires some local information. Note that the FC can reasonably observe the pilot signals $\{p_1[t]\}$ and $\{p_2[t]\}$ in the same way SUs do. Then, the FC needs to know the local random variables $\{V_\tau^{ikn}\}_{i,k,n}$ at the stopping time τ . In a straightforward way SUs can quantize and send $\{V_\tau^{ikn}\}_{i,k,n}$ at time τ . However, this method has several disadvantages in practice. Firstly, it needs high bandwidth at time τ on each reporting channel between SUs and the FC. Moreover, the reporting channels are utilized inefficiently. They remain idle until time τ , and at time τ each SU sends a number of bits, which may cause congestion at the FC. To overcome these practical issues SUs can sequentially report $\{V_\tau^{ikn}\}_{i,k,n}$. For sequential reporting level-triggered sampling, a non-uniform sampling technique, was shown to be much superior to the traditional uniform sampling in terms of bandwidth and energy requirements for detection and estimation purposes in [22] and [23], respectively. Therefore, we propose that SUs sequentially report $\{V_\tau^{ikn}\}_{i,k,n}$ using level-triggered sampling.

B. Level-triggered Sampling

Each SU k , via the same level-triggered sampling procedure, informs the FC whenever considerable change occurs in its four local processes $\{V_t^{ikn}\}$, $i = 1, 2$, $n = 1, 2$. In other words, $4K$ identical samplers run in parallel for $4K$ different processes. Hence, we will describe the procedure for a single

process $\{V_t^{ikn}\}_t$. The level-triggered sampling is a simple form of event-triggered sampling, in which sampling (communication) times $\{t_m\}$, $m \in \mathbb{N}$, are not deterministic, but rather dynamically determined by the random process $\{V_t^{ikn}\}_t$, i.e.,

$$t_m \triangleq \min\{t > t_{m-1} : V_t^{ikn} - V_{t_{m-1}}^{ikn} \notin (-\Delta, \Delta)\}, \quad m \in \mathbb{N}, \quad t_0 = 0. \quad (20)$$

The threshold parameter Δ is a constant known by both SUs and the FC. At each sampling time t_m , SU k transmits r bits, $b_{m,1}b_{m,2} \dots b_{m,r}$, to the FC. The first bit, $b_{m,1}$, indicates the threshold crossed (either Δ or $-\Delta$) by the incremental process $v_m \triangleq V_{t_m}^{ikn} - V_{t_{m-1}}^{ikn}$, i.e.,

$$b_{m,1} = \text{sign}(v_m). \quad (21)$$

The remaining $r - 1$ bits are used to quantize the over(under)shoot $q_m \triangleq |v_m| - \Delta$ into \tilde{q}_m . At each sampling time t_m , the overshoot value q_m cannot exceed the magnitude of the last sample $|y_{ik}^n[t_m]|$ in the incremental process $v_m = \sum_{t=t_{m-1}+1}^{t_m} y_{ik}^n[t]$. The quantization interval $[0, \phi]$ is uniformly divided into 2^{r-1} subintervals with the step size $\frac{\phi}{2^{r-1}}$. The mid value of each subinterval is used as the corresponding quantization level, i.e., a mid-riser quantizer is used. When $q_m > \phi$, the uppermost quantization level is used. The parameter ϕ is determined so that $P(q_m > \phi)$ is sufficiently small. From [22, Section IV-B] we can set the threshold Δ using

$$\Delta \tanh\left(\frac{\Delta}{2}\right) = \frac{1}{M} \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 |\mathbb{E}_i[V_1^{ikn}]| \quad (22)$$

for the FC to receive messages with an average rate of M messages per unit time under H_i , $i = 0, 1$. In Fig. 4, the level-triggered sampling procedure is demonstrated on a sample path of V_t^{ikn} .

The FC, upon receiving the bits $b_{m,1}b_{m,2} \dots b_{m,r}$ from SU k at time t_m , recovers the quantized value of v_m by computing

$$\tilde{v}_m \triangleq b_{m,1}(\Delta + \tilde{q}_m). \quad (23)$$

Then, it sequentially sums up $\{\tilde{v}_m\}$, at the sampling (communication) times $\{t_m\}$ to obtain an approximation \tilde{V}_t^{ikn} to the sufficient statistic V_t^{ikn} , i.e.,

$$\tilde{V}_t^{ikn} \triangleq \sum_{m=1}^{M_t} \tilde{v}_m, \quad (24)$$

where M_t is the number of messages that the FC receives from SU k about the process $\{V_t^{ikn}\}$ up to time t . During the times the FC receives no message, i.e., $t \notin \{t_m\}$, \tilde{V}_t^{ikn} is kept constant.

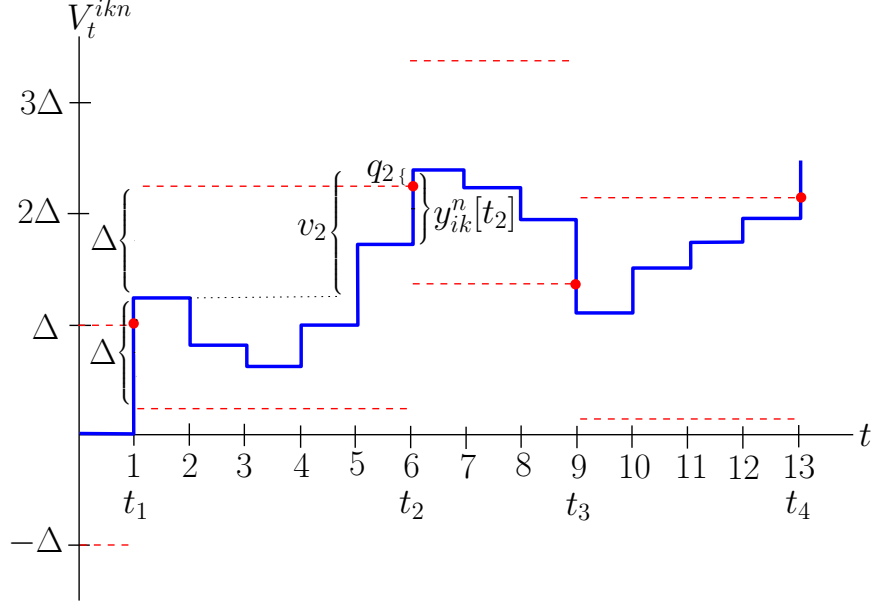


Fig. 4. The level-triggered sampling procedure used at SUs.

At the stopping time τ , given by (13), the FC estimates each channel coefficient h_{ik}^n using

$$\tilde{h}_{ik}^n[\tau] = \frac{\tilde{V}_t^{ikn} + \frac{\mu_{ik}}{2} \frac{N_0^{ik}}{\sigma_{ik}^2}}{U_t^i + \frac{N_0^{ik}}{\sigma_{ik}^2}}, \quad (25)$$

and decides according to the following rule

$$\tilde{d}_\tau = \begin{cases} 1 & \text{if } \tilde{L}_\tau \geq \log \frac{c_0}{c_1 + c_e \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 (\tilde{h}_{ik}^n[\tau])^2}, \\ 0 & \text{otherwise} \end{cases}, \quad (26)$$

where $\tilde{L}_\tau = \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 \tilde{L}_\tau^{ikn}$, and \tilde{L}_τ^{ikn} is computed from (11) by substituting \tilde{V}_t^{ikn} for V_t^{ikn} .

After making a decision, the FC grants the transmission privilege to the SU Tx with the highest achievable throughput. When the decision is in favor of H_0 , i.e., $\tilde{d}_\tau = 0$, one of them is selected randomly (or in some specific order) since in this case any SU Tx can transmit with its maximum power P_{\max} . On the other hand, when $\tilde{d}_\tau = 1$, the FC selects SU Tx k^* where

$$k^* \triangleq \arg \max_{k_t} \left\{ \min \left\{ \frac{I_1}{|\tilde{h}_{1k_t}[\tau]|^2}, \frac{I_2}{|\tilde{h}_{2k_t}[\tau]|^2} \right\} \right\}, \quad (27)$$

k_t is the SU Tx index, and $|\tilde{h}_{ik}[\tau]|^2 = \sum_{n=1}^2 (\tilde{h}_{ik}^n[\tau])^2$, $i = 1, 2$. The pseudocodes for the procedures at SU k and the FC in the proposed SJDE-based dynamic spectrum access method (DSA-SJDE) are given in Algorithms 1 and 2, respectively. In Algorithm 1, an SU Rx never executes lines 17-23 since the FC reports d_τ to SU Tx k^* (cf. line 20 in Algorithm 2).

Algorithm 1 DSA-SJDE procedure at SU k

```

1: Initialization:  $\{t, m_{in}, v_{in}, V_{in}, U_i\} \leftarrow 0, \forall i, n$ 
2: while  $|v_{in}| < \Delta, \forall i, n$  and  $\sum_{i=1}^2 U_i < \gamma$  do
3:    $t \leftarrow t + 1$ 
4:    $v_{in} \leftarrow v_{in} + y_{ik}^n[t]$ 
5:    $V_{in} \leftarrow V_{in} + y_{ik}^n[t]$ 
6:    $U_i \leftarrow U_i + |p_i[t]|^2$ 
7: end while
8: if  $|v_{in}| \geq \Delta$  {for any  $i, n$ } then
9:    $m_{in} \leftarrow m_{in} + 1$ 
10:   $t_m^{in} = t$ 
11:  Send  $b_{m,1}^{in} = \text{sign}(v_{in})$  and  $r - 1$  quantization bits for  $q_m^{in} = v_{in} - \Delta$  to FC
12:   $v_{in} \leftarrow 0$ 
13: end if
14: if  $\sum_{i=1}^2 U_i \geq \gamma$  or  $t \geq T_p$  then
15:    $\tau = t$ 
16:   if FC reports  $d_\tau$  then
17:     if  $d_\tau = 0$  then
18:        $P = P_{\max}$ 
19:     else
20:       Compute  $\hat{h}_{in}$  as in (15) using  $V_{in}$  and  $U_i$ 
21:        $P = \min \left\{ P_{\max}, \frac{\hat{I}_1^\tau}{\sum_{n=1}^2 (h_{1n})^2}, \frac{\hat{I}_2^\tau}{\sum_{n=1}^2 (h_{2n})^2} \right\}, \{ \text{see (29) for } \hat{I}_i^\tau \}$ 
22:     end if
23:     Start data transmission with power  $P$ 
24:   else
25:     Stop
26:   end if
27: else
28:   Go to line 2
29: end if

```

C. Discussions

The procedures at SUs and the FC, given in Algorithms 1 and 2, restarts at the beginning of each frame with duration T (see Fig. 2). Each SU k performs the procedure in Algorithm 1. The stopping threshold γ is selected through offline simulations to maximize the average SU throughput in DSA-SJDE, given by

$$\bar{R} = \mathbb{E} \left[\frac{T - \tau}{T} \left\{ [\pi_0(1 - P_f) + (1 - \pi_0)P_m]\Gamma_0 + [\pi_0P_f + (1 - \pi_0)(1 - P_m)]\Gamma_1 \right\} \right] \quad (28)$$

where $\Gamma_0 \triangleq \log \left(1 + \frac{|\beta_{k^*}|^2 P_{\max}}{N_0^{k^*}} \right)$, $\Gamma_1 \triangleq \log \left(1 + \frac{|\beta_{k^*}|^2 P_{k^*}^\tau}{N_0^{k^*} + |h_{1k_r^*}|^2 Q_1 + |h_{2k_r^*}|^2 Q_2} \right)$, k_r^* denotes the SU Rx corresponding to SU Tx k^* , π_0 is the prior probability for the hypothesis H_0 , P_f is the false alarm probability, i.e., $P_0(\tilde{d}_\tau = 1)$, and P_m is the misdetection probability, i.e., $P_1(\tilde{d}_\tau = 0)$. The sensing time τ is governed

Algorithm 2 DSA-SJDE procedure at FC

```

1: Initialization:  $\{V_{ikn}, U_i\} \leftarrow 0, \forall i, k, n$ 
2: while  $\sum_{i=1}^2 U_i < \gamma$  or  $t < T_p$  do
3:    $t \leftarrow t + 1$ 
4:    $U_i \leftarrow U_i + |p_i[t]|^2$ 
5:   if  $b_{m,1}^{ikn} \dots b_{m,r}^{ikn}$  received {for any  $i, k, n$ } then
6:     Compute  $q_m^{ikn}$  from  $b_{m,2}^{ikn} \dots b_{m,r}^{ikn}$ 
7:      $V_{ikn} = V_{ikn} + b_{m,1}^{ikn}(\Delta + q_m^{ikn})$ 
8:   end if
9: end while
10:  $\tau = t$ 
11: Compute  $\tilde{h}_{ikn}$  from (25) using  $V_{ikn}$  and  $U_i, \forall i, k, n$ 
12: Compute  $L_{ikn}$  from (11) using  $V_{ikn}$  and  $U_i, \forall i, k, n$ 
13:  $L = \sum_{k=1}^K \sum_{i=1}^2 \sum_{n=1}^2 L_{ikn}$ 
14: Compute  $d_\tau$  from (26) using  $L$  and  $\{\tilde{h}_{ikn}\}$ 
15: if  $d_\tau = 0$  then
16:   Select  $k^*$  randomly or in some specific order from SU transmitters
17: else
18:   Find  $k^*$  as in (27) using  $\{\tilde{h}_{ikn}\}$ 
19: end if
20: Report  $d_\tau$  to SU Tx  $k^*$ , and instruct the others to stop

```

by the threshold γ . The scaling term $\frac{T-\tau}{T}$ in (28) represents the throughput penalty due to sensing. Hence, small threshold γ on average increases the scaling term, affecting \bar{R} positively. On the other hand, it causes larger error probabilities, P_f and P_m . Note that $P_{k^*}^\tau = \min \left\{ P_{\max}, \frac{\hat{I}_{1k^*}^\tau}{|h_{1k^*}[\tau]|^2}, \frac{\hat{I}_{2k^*}^\tau}{|h_{2k^*}[\tau]|^2} \right\} \leq P_{\max}$, thus $\Gamma_0 > \Gamma_1$. As a result, increasing P_f decreases \bar{R} . Although it looks like \bar{R} is directly proportional to P_m , large P_m values are not feasible due to the interference constraints. This defines a lower bound on the stopping threshold γ . As clearly seen, there is a trade-off in selecting the γ value. It is convenient to find the best γ value, that maximizes \bar{R} , performing an offline numerical search in the interval $[\gamma_0, \gamma_1]$. The lower bound γ_0 is determined by the interference constraints as mentioned earlier. We need the upper bound γ_1 to control the probability that the sensing time exceeds the preamble duration, i.e., $P(\tau > T_p)$, where the signal model in (4) is valid. In such an exceptional case, when $\tau > T_p$, the sensing and estimation should terminate, i.e., $\tau = T_p$, since the signal model is no more valid.

When PU communication is detected, i.e., $\tilde{d}_\tau = 1$, the SU selected for data transmission needs to use calibrated maximum interference levels $\hat{I}_{ik^*}^\tau \triangleq \alpha_{ik^*}^\tau I_i$, instead of original values I_i , $i = 1, 2$, in computing its transmission power. This is required to compensate for estimation errors. To satisfy the

interference constraints we should have

$$\frac{\alpha_{ik^*}^\tau I_i}{|\hat{h}_{ik^*}[\tau]|^2} |h_{ik^*}|^2 \leq I_i, \quad i = 1, 2, \quad (29)$$

hence $\alpha_{ik^*}^\tau \leq \frac{|\hat{h}_{ik^*}[\tau]|^2}{|h_{ik^*}|^2}$ with a high probability. Since the actual channel coefficient h_{ik^*} is unknown, through offline simulations we set $\alpha_{ik_t}^\tau$ for each $\tau \in (0, T_p]$ so that $P\left(\frac{|\hat{h}_{ik_t}[\tau]|^2}{|h_{ik_t}|^2} \geq \alpha_{ik_t}^\tau\right)$ is sufficiently high. Note that there are two sources that cause excess interference over I_i , namely misdetection and the event $\frac{|\hat{h}_{ik^*}[\tau]|^2}{|h_{ik^*}|^2} < \alpha_{ik^*}^\tau$. The probabilities $P_m = P_1(\tilde{d}_\tau = 0)$ and $P\left(\frac{|\hat{h}_{ik^*}[\tau]|^2}{|h_{ik^*}|^2} < \alpha_{ik^*}^\tau\right)$ should be made sufficiently small in order to meet the PU outage constraints.

V. SIMULATION RESULTS

In this section, we provide simulation results to compare different spectrum access methods in terms of the average SU throughput. We first consider two conventional methods: underlay and opportunistic access. These two methods have intrinsic deficiencies. In the former the SU is blind to the idle state of PUs, and in the latter it is unable to benefit from deep fades in cross links. It could be anticipated that a combination of these two methods, as in DSA-SJDE and DSA-SPRT, may result in a higher SU throughput. DSA-SPRT is the straightforward sequential implementation of such combination. It uses SPRT for spectrum sensing, MMSE estimator for channel estimation, and uniform sampling for distributed operation. On the other hand, DSA-SJDE, the proposed novel spectrum access method, uses the SJDE for sensing and estimation, and level-triggered sampling for distributed implementation. In the opportunistic access scheme, we use the LRT for sensing and the traditional uniform sampling for distributed implementation. In the underlay scheme, we assume that SUs somehow perfectly estimate the channel coefficients during the preamble.

We plot the average SU throughput \bar{R} against the outage probability constraint P_{out} , the maximum transmission power P_{max} for SU, the prior probability π_0 of idle PU, and the fraction $\frac{T}{T_p}$ of frame length to the preamble duration respectively in the subsequent figures. The preamble duration is fixed at $T_p = 10$ ms and the global clock runs, i.e., PUs transmit pilot symbols and SUs observe discrete-time samples, with a frequency of $f_s = 1$ MHz. In PU communication 16-QAM is used with an average power $E[|p_i[t]|^2] = P_i = 1$. PUs utilize random number generators, whose seeds are known to SUs and the FC, to generate pilot symbols in the preamble. All simulated channels are Rayleigh fading channels, i.e., channel coefficient h_{ik} is proper complex Gaussian random variable with zero mean and finite variance σ_{ik}^2 . We set $N_0^{ik} = \sigma_{ik}^2 = 1$, hence $\text{SNR} = E[|p_i[t]|^2] = 1$ (0 dB) under H_1 . In opportunistic access and DSA-SPRT, the period of uniform sampling for reporting V_t^{ikn} is set as four unit time, i.e., $T_u = 4T_s = \frac{4}{f_s}$. Since

each SU samples four processes, the FC receives K messages per unit time, T_s . For a fair comparison we set the average message rate of level-triggered sampling to the same value, i.e., $M = K$. Then, using (22) the corresponding value of the sampling threshold Δ is found. Throughout this section we simulate a two-SU system, i.e., $K = 2$.

We use a 50% safety margin while determining the maximum interference level I_i from P_{out} using (2). Moreover, as additional safety measures to protect the PU QoS, i.e., to satisfy the P_{out} constraint, we determine α_{i1}^τ as the fifth percentile of $\frac{|\hat{h}_{i1}[\tau]|^2}{|h_{i1}|^2}$ to calibrate the maximum interference levels at SUs, and confine the misdetection probability P_m to values smaller than $P_{\text{out}}/5$. For the DSA-SPRT, DSA-SJDE, and the opportunistic access scheme, through offline simulations we find the best parameters that maximize \bar{R} , complying with the constraint $P_m < P_{\text{out}}/5$. Specifically, via offline numerical search, we use the optimum values for the threshold pair in SPRT, the stopping threshold γ in SJDE, the deterministic sensing time τ and the LRT threshold in the opportunistic access scheme. We use $c_0 = c_1 = 0.2$, $c_e = 0.6$ for SJDE as in Section III.

SU throughput vs. PU outage probability: In the first set of simulations, we set $P_{\text{max}} = 15$ dB, $\pi_0 = 0.5$, $T = 10 \times T_p$, and vary $P_{\text{out}} \in [0.025, 0.125]$. In this case, the maximum interference levels I_i vary between -9 dB and 6 dB.

In Fig. 5, we see that the proposed spectrum access schemes with sequential detectors and estimators, being combinations of conventional methods, perform better than the underlay and the opportunistic access schemes, as expected. Not surprisingly, the underlay scheme performs poorly under strict outage probability (interference) constraints, and considerably improves its performance as the constraints relax because its transmit power solely depends on the maximum interference levels. Conversely, the opportunistic access scheme is mostly unaffected by the changing outage probability constraint as it does not utilize the maximum interference levels to determine its transmit power. The slight performance increase as P_{out} grows is due to the relaxation on the P_m constraint. On the other hand, the sequential schemes, being combinations of the conventional approaches, enjoy the advantages of opportunistic access and underlay when P_{out} is small and large, respectively. Moreover, the novel DSA-SJDE scheme significantly outperforms DSA-SPRT, which uses well-known techniques for sampling and distributed implementation, due to its distinct features: the joint nature of detector and estimator (cf. Section III), the separation property of stopping rule and detector (cf. Section III), and the adaptive nature of level-triggered sampling (cf. Section IV-A). Note that the estimator provides some side information about the true hypothesis, and thus its incorporation into the decision function improves the SU throughput, which is a joint function of detector and estimator. For the advantages of the latter two features we refer to Section III and Section

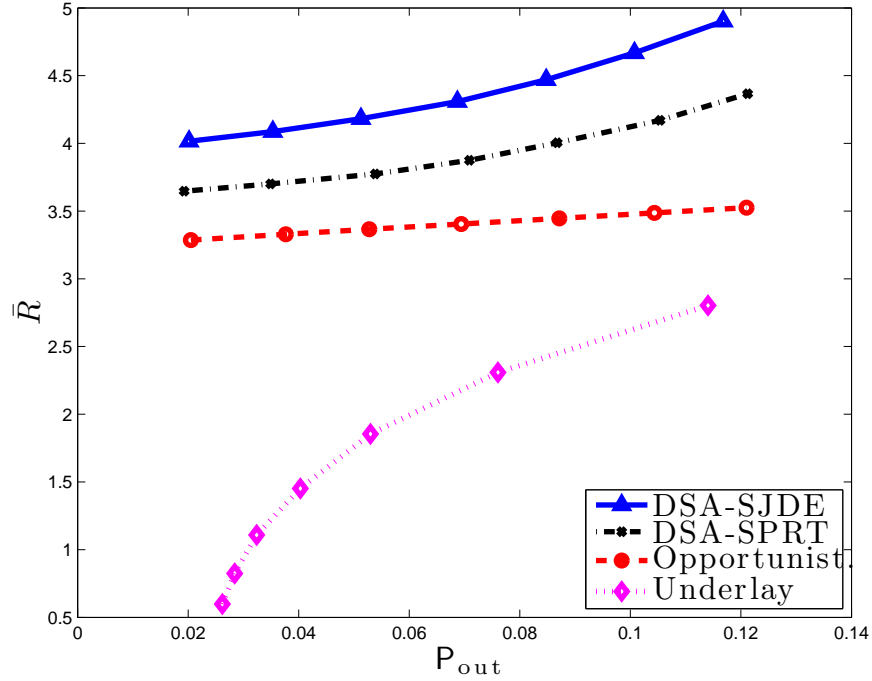


Fig. 5. Average SU throughput vs. PU outage probability for the conventional (underlay and opportunistic) and the proposed (DSA-SJDE and DSA-SPRT) dynamic spectrum access schemes.

IV-A, respectively.

SU throughput vs. SU maximum power: We next plot \bar{R} vs. $P_{max} \in [9 \text{ dB}, 21 \text{ dB}]$ in Fig. 6, where $P_{out} = 0.075$, $\pi_0 = 0.5$, $T = 10 \times T_p$. In this figure, we see that the sensing-based-schemes greatly benefit from increasing P_{max} as they set their transmit power to P_{max} when H_0 is decided. In contrast, in the underlay scheme, where no spectrum sensing is performed, the direct effect of increasing P_{max} is not observed. For small P_{max} values, the utility of spectrum sensing is deemphasized, and the advantage of the perfect CSI assumption of the underlay scheme becomes apparent. It is again notable that the proposed sequential schemes, especially DSA-SJDE, considerably outperform the conventional methods.

SU throughput vs. H_0 prior probability: In the next set of simulations, we investigate the effect of the prior probability π_0 of H_0 on the average SU throughput, \bar{R} , while we set $P_{max} = 15 \text{ dB}$, $P_{out} = 0.075$, and $T = 10 \times T_p$. Because of the same reason in the changing P_{max} case the sensing-based-schemes significantly improve their performances with increasing π_0 , as shown in Fig. 7. The advantage of perfect CSI in the underlay scheme is even more emphasized here, e.g., underlay outperforms the sensing-based-schemes for $\pi_0 = 0$. The slight improvement in the underlay performance with increasing π_0 is due to the lack of interference at the SU receiver under H_0 .

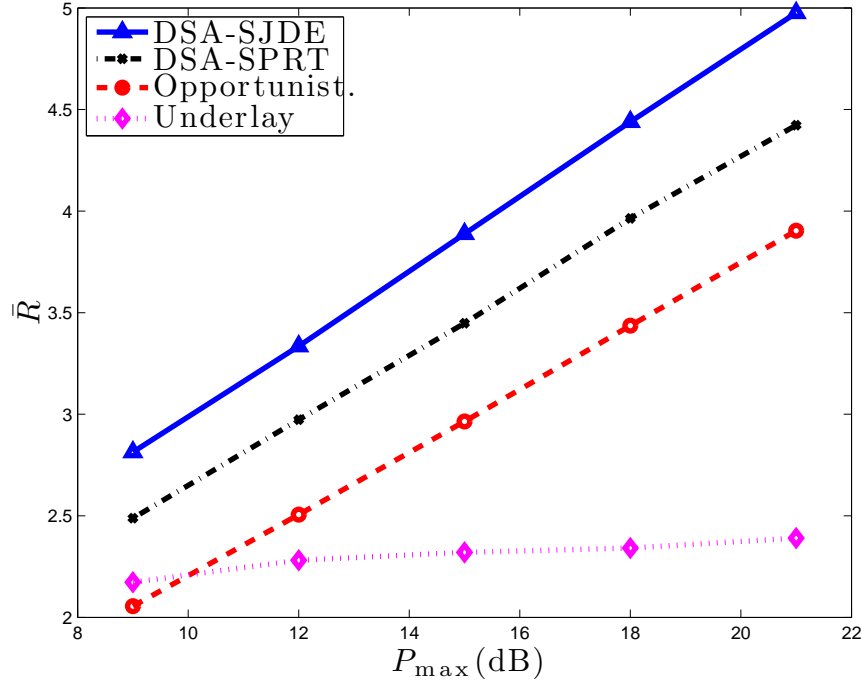


Fig. 6. Average SU throughput vs. SU maximum power for the conventional (underlay and opportunistic) and the proposed (DSA-SJDE and DSA-SPRT) dynamic spectrum access schemes.

SU throughput vs. Frame length: Finally, in Fig. 8 we set $P_{\max} = 15$ dB, $P_{\text{out}} = 0.075$, $\pi_0 = 0.5$, and analyze the effect of the frame length T on \bar{R} . Note that T corresponds to the coherence time in the system. PUs carry out preamble communication every T seconds to estimate the changing channel coefficients. In other words, it is assumed that the channels do not change during each frame of length T . Hence, small T corresponds to fast fading channels, whereas large T implies slow fading channels.

Changing T while keeping the preamble duration T_p fixed does not affect the detection and estimation performances, but only changes the remaining time for data transmission, i.e., the scaling term in the \bar{R} expression in (28). Since the scaling term is common to all schemes, they all exhibit similar behaviors with changing T . After some certain value, e.g., $T/T_p = 10$, the scaling term well approximates unity, and as a result the throughput curves saturate.

VI. CONCLUSIONS

We have considered dynamic spectrum access under statistical CSI. For a cognitive radio network, a cooperative scheme based on sequential joint spectrum sensing and channel estimation has been proposed. With the objective of SU throughput maximization subject to PU outage constraints, the sensing time

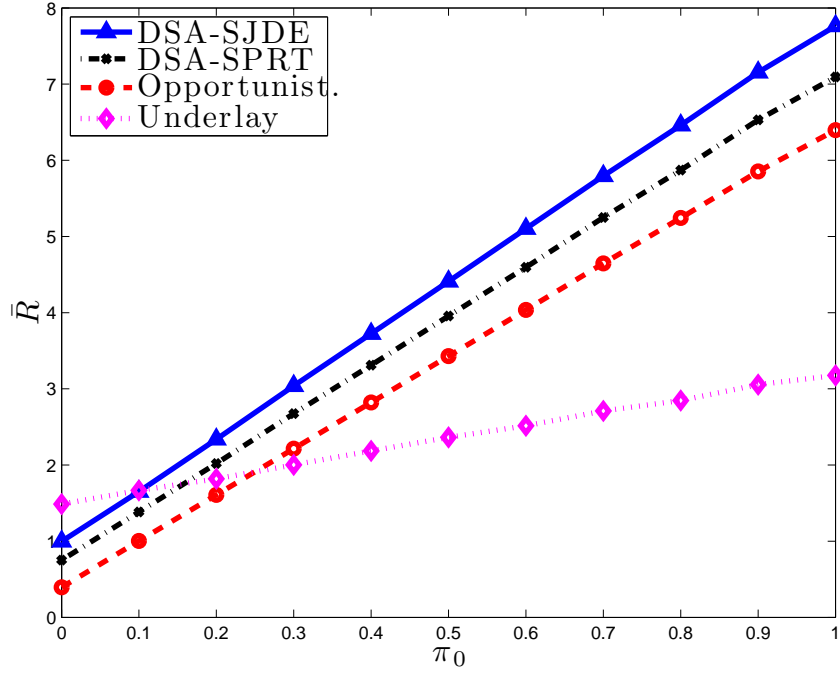


Fig. 7. Average SU throughput vs. H_0 prior probability for the conventional (underlay and opportunistic) and the proposed (DSA-SJDE and DSA-SPRT) dynamic spectrum access schemes.

needs to be minimized, hence the sequential framework is a better fit to the problem of interest than the fixed-sample-size framework. Unlike the existing works in the literature, channel estimation, which is of practical interest, has been included in the problem formulation. A salient feature of the proposed scheme is that the sensing decision rule makes use of the side information on the true hypothesis provided by the channel estimator. A bandwidth and energy-efficient nonuniform sampling technique, called level-triggered sampling, is used to transmit the information from SUs to the FC, which makes the spectrum sensing decision. Then, the sensing decision and the channel estimates are employed to determine the SU transmit power. Through simulations we have shown the superior performance of the proposed scheme in terms of the average SU throughput over its counterpart that treat the sensing and estimation problems separately, and the conventional spectrum access methods (underlay and opportunistic access) under the same PU outage constraints.

APPENDIX: PROOF OF THEOREM 2

As shown in [13] the optimum estimators, decision function, and the stopping time can be found separately, i.e., we can fix two of them, and find the optimum solution for the remaining one. Furthermore,

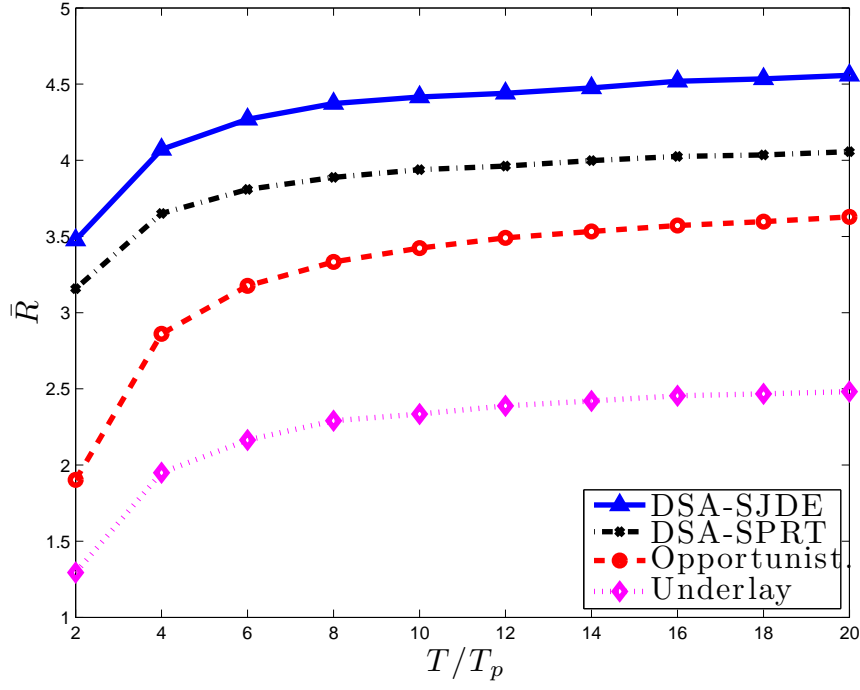


Fig. 8. Average SU throughput vs. frame length in terms of preamble duration for the conventional (underlay and opportunistic) and the proposed (DSA-SJDE and DSA-SPRT) dynamic spectrum access schemes.

since $\{h_i^n\}$ are independent, we can minimize each MSE term individually over the corresponding estimator. Hence, the MMSE estimator in (15) is the optimum estimator for each h_i^n .

Next, substituting the MMSE estimates of $\{h_i^n\}$ into (12) we seek the optimum decision rule. From the classical estimation theory (e.g., [29, page 151]) we know that the conditional mean of the parameter to be estimated gives the MMSE estimator, i.e., $\bar{E}_1[h_i^n|\mathcal{G}_t] = \hat{h}_i^n[t]$, and its conditional variance is $\bar{E}_1\left[\left(h_i^n - \hat{h}_i^n[t]\right)^2\right] = \frac{N_0^i/2}{U_t^i + \frac{N_0^i}{\sigma_i^2}}$. Hence, using

$$\begin{aligned}
 \bar{E}_1\left[\left(\hat{h}_i^n[\tau] - h_i^n\right)^2 \mathbb{1}_{\{d_\tau=1\}}|\mathcal{F}_\tau\right] &= \sum_{t=0}^{\infty} \bar{E}_1\left[\frac{N_0^i/2}{U_t^i + \frac{N_0^i}{\sigma_i^2}} \mathbb{1}_{\{d_t=1\}}|\mathcal{F}_t\right] \mathbb{1}_{\{\tau=t\}} \\
 &= \sum_{t=0}^{\infty} \frac{N_0^i/2}{U_t^i + \frac{N_0^i}{\sigma_i^2}} P_1(d_t = 1|\mathcal{F}_t) \mathbb{1}_{\{\tau=t\}} = \frac{N_0^i/2}{U_\tau^i + \frac{N_0^i}{\sigma_i^2}} P_1(d_\tau = 1|\mathcal{F}_\tau) \quad (30)
 \end{aligned}$$

and

$$\begin{aligned}
\bar{\mathbb{E}}_1 \left[(h_i^n)^2 \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right] &= \sum_{t=0}^{\infty} \bar{\mathbb{E}}_1 \left[(h_i^n)^2 \mathbb{1}_{\{d_t=0\}} | \mathcal{F}_t \right] \mathbb{1}_{\{\tau=t\}} \\
&= \sum_{t=0}^{\infty} \mathbb{E}_1 \left[\bar{\mathbb{E}}_1 \left[(h_i^n)^2 | \mathcal{G}_t \right] \mathbb{1}_{\{d_t=0\}} | \mathcal{F}_t \right] \mathbb{1}_{\{\tau=t\}} = \mathbb{E}_1 \left[\left(\hat{h}_i^n[\tau] \right)^2 \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right] + \frac{N_0^i/2}{U_\tau^i + \frac{N_0^i}{\sigma_i^2}} \mathbb{P}_1(d_\tau = 0 | \mathcal{F}_\tau)
\end{aligned} \tag{31}$$

we can rewrite the cost in (12) as

$$\begin{aligned}
\mathcal{C}(\tau, d_\tau) &= c_0 \mathbb{P}_0(d_\tau = 1 | \mathcal{F}_\tau) + c_1 \mathbb{P}_1(d_\tau = 0 | \mathcal{F}_\tau) \\
&\quad + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\mathbb{E}_1 \left[\left(\hat{h}_i^n[\tau] \right)^2 \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right] + \frac{N_0^i/2}{U_\tau^i + \frac{N_0^i}{\sigma_i^2}} \right), \tag{32}
\end{aligned}$$

where $\hat{h}_i^n[\tau]$ is given by (15). Since the last term in (32) does not depend on d_τ , we consider only the remaining terms, i.e.,

$$\tilde{\mathcal{C}}(\tau, d_\tau) = c_0 \mathbb{P}_0(d_\tau = 1 | \mathcal{F}_\tau) + c_1 \mathbb{P}_1(d_\tau = 0 | \mathcal{F}_\tau) + c_e \sum_{i=1}^2 \sum_{n=1}^2 \mathbb{E}_1 \left[\left(\hat{h}_i^n[\tau] \right)^2 \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right]. \tag{33}$$

We next combine the terms on the right-hand side of (33) under \mathbb{E}_0 by changing the measure under \mathbb{H}_1 to its counterpart under \mathbb{H}_0 . The likelihood ratio $\frac{f_1(\{y_i^n[t], p_i[t]\})}{f_0(\{y_i^n[t], p_i[t]\})} = e^{L_\tau}$ is used for change of measures.

$$\begin{aligned}
\tilde{\mathcal{C}}(\tau, d_\tau) &= \mathbb{E}_0 \left[c_0 \mathbb{1}_{\{d_\tau=1\}} + e^{L_\tau} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[\tau] \right)^2 \right\} \mathbb{1}_{\{d_\tau=0\}} | \mathcal{F}_\tau \right] \\
&= \sum_{t=0}^{\infty} \mathbb{E}_0 \left[c_0 \mathbb{1}_{\{d_t=1\}} + e^{L_t} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[t] \right)^2 \right\} \mathbb{1}_{\{d_t=0\}} | \mathcal{F}_t \right] \mathbb{1}_{\{\tau=t\}} \\
&= \sum_{t=0}^{\infty} \mathbb{E}_0 \left[\left(c_0 - e^{L_t} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[t] \right)^2 \right\} \right) \mathbb{1}_{\{d_t=1\}} | \mathcal{F}_t \right] \mathbb{1}_{\{\tau=t\}} \\
&\quad + \sum_{t=0}^{\infty} \mathbb{E}_0 \left[e^{L_t} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[t] \right)^2 \right\} | \mathcal{F}_t \right] \mathbb{1}_{\{\tau=t\}}, \tag{34}
\end{aligned}$$

The optimum decision rule that minimizes (34) selects \mathbb{H}_1 , i.e., $d_\tau = 1$, when

$$c_0 \leq e^{L_t} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[\tau] \right)^2 \right\},$$

and selects \mathbb{H}_0 otherwise, proving (14).

Finally, substituting the optimum detector into the cost function (32) we have

$$\mathcal{C}(\tau) = \mathbb{E}_0 \left[\left(c_0 - e^{L_\tau} \left\{ c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\hat{h}_i^n[\tau] \right)^2 \right\} \right)^- \middle| \mathcal{F}_\tau \right] + c_1 + c_e \sum_{i=1}^2 \sum_{n=1}^2 \left(\mathbb{E}_1 \left[\left(\hat{h}_i^n[\tau] \right)^2 \middle| \mathcal{F}_\tau \right] + \frac{N_0^i/2}{U_\tau^i + \frac{N_0^i}{\sigma_i^2}} \right), \quad (35)$$

where $(x)^- = \min(x, 0)$ is the negative part operator. We now focus on $\mathbb{E}_1 \left[\left(\hat{h}_i^n[t] \right)^2 \middle| \mathcal{F}_\tau \right]$, where

$$\hat{h}_i^n[t] = \frac{V_t^{in} + \frac{\mu_i}{2} \frac{N_0^i}{\sigma_i^2}}{U_t^i + \frac{N_0^i}{\sigma_i^2}} \text{ is given by (15). Note from (4) that under } H_1 \text{ given } \mathcal{F}_t \text{ we have } V_t^{in} = \sum_{m=1}^t y_i^n[m] \sim \mathcal{N} \left(\frac{\mu_i}{2} U_t^i, \frac{\sigma_i^2}{2} \sum_{m=1}^t |p_i[m]|^4 + \frac{N_0^i}{2} U_t^i \right), \text{ hence } \hat{h}_i^n[t] \text{ is Gaussian with mean } \frac{\mu_i}{2} \text{ and variance } \frac{\frac{\sigma_i^2}{2} \sum_{m=1}^t |p_i[m]|^4 + \frac{N_0^i}{2} U_t^i}{\left(U_t^i + \frac{N_0^i}{\sigma_i^2} \right)^2}.$$

Therefore, $\mathbb{E}_1 \left[\left(\hat{h}_i^n[t] \right)^2 \middle| \mathcal{F}_\tau \right] = \frac{\mu_i^2}{4} + \frac{\frac{\sigma_i^2}{2} \sum_{m=1}^t |p_i[m]|^4 + \frac{N_0^i}{2} U_t^i}{\left(U_t^i + \frac{N_0^i}{\sigma_i^2} \right)^2}$, which is decreasing in U_t^i . As a result, the last term in (35) is decreasing in U_t . Indeed the first term is also decreasing in U_t , hence the optimum stopping rule is a thresholding on the conditional Fisher information U_t as shown in (13). The analysis of the first term, which is very technical and involved, directly follows from [13, Theorem 1], thus is omitted here.

REFERENCES

- [1] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access: Signal processing, networking, and regulatory policy," *IEEE Signal Processing Mag.*, vol.24, no. 3, pp. 79–89, May 2007.
- [2] X. Kang, Y. C. Liang, A. Nallanathan, H. K. Garg, and R. Zhang, "Optimal power allocation for fading channels in cognitive radio networks: ergodic capacity and outage capacity," *IEEE Trans. Wireless Commun.*, vol. 8, no.2, pp. 940–950, Feb. 2009.
- [3] L. Musavian and S. Aissa, "Capacity and power allocation for spectrum sharing communications in fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no.1, pp. 148–156, 2009.
- [4] X. Kang, R. Zhang, Y. C. Liang, and H. K. Garg, "Optimal power allocation strategies for fading cognitive radio channels with primary user outage constraint," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 374–383, 2011.
- [5] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol.7, no. 4, pp. 1326–1337, Apr. 2008.
- [6] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 2053–2071, May 2008.
- [7] R. Fan, H. Jiang, Q. Guo, and Z. Zhang, "Joint optimal cooperative sensing and resource allocation in multichannel cognitive radio networks," *IEEE Tran. Veh. Technol.*, vol. 60, no. 2, pp. 722–729, Feb. 2011.
- [8] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

- [9] X. Kang, Y. C. Liang, H. K. Garg, and L. Zhang, "Sensing-based spectrum sharing in cognitive radio networks", *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4649–4654, Oct. 2009.
- [10] R. Zhang, "On peak versus average interference power constraints for protecting primary users in cognitive radio networks", *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 2112–2120, Apr. 2009.
- [11] Z. Chen, X. Wang, and X. Zhang, "Continuous power allocation strategies for sensing-based multiband spectrum sharing," *IEEE J. Sel. Areas Commun.*, to be published.
- [12] G. V. Moustakides, G. H. Jajamovich, A. Tajer, and X. Wang, "Joint detection and estimation: optimum tests and applications," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4215–4229, July 2012.
- [13] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Sequential joint detection and estimation," *SIAM Theory Probab. Appl.*, to be published, <http://arxiv.org/abs/1302.6058>.
- [14] D. Middleton and R. Esposito, "Simultaneous optimum detection and estimation of signals in noise," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 3, pp. 434–444, May 1968.
- [15] A. Fredriksen, D. Middleton, and D. Vandelinde, "Simultaneous signal detection and estimation under multiple hypotheses," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 5, pp. 607–614, Sep. 1972.
- [16] G. Yue, X. Zhou, and X. Wang, "Performance comparisons of channel estimation techniques in multipath fading CDMA," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 716–724, May 2004.
- [17] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, July 2000.
- [18] H. Tang, "Some physical layer issues of wide-band cognitive radio systems," in *Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, Maryland, USA, pp. 151–159, Nov. 2005.
- [19] A. Sahai, R. Tandra, S. M. Mishra, and N. Hoven, "Fundamental design tradeoffs in cognitive radio systems," in *Proc. of Int. Workshop on Technology and Policy for Accessing Spectrum*, Aug. 2006.
- [20] S. M. Mishra, S. t. Brink, R. Mahadevappa, and R. W. Brodersen, "Cognitive technology for ultra-wideband/WiMax coexistence," in *Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Dublin, Ireland, pp. 179–186, Apr. 2007.
- [21] G. Fellouris and G.V. Moustakides, "Decentralized sequential hypothesis testing using asynchronous communication," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 534–548, Jan. 2011.
- [22] Y. Yilmaz, G.V. Moustakides, and X. Wang, "Cooperative sequential spectrum sensing based on level-triggered sampling," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4509–4524, Sep. 2012.
- [23] Y. Yilmaz, and X. Wang, "Sequential Decentralized Parameter Estimation under Randomly Observed Fisher Information," *IEEE Trans. Inf. Theory*, to be published, <http://arxiv.org/abs/1211.3720>.
- [24] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Stat.*, vol. 19, pp. 326–329, 1948.
- [25] B.K. Ghosh, "On the attainment of the Cramer-Rao bound in the sequential case," *Sequential Analysis*, vol. 6, no. 3, pp. 267–288, 1987.
- [26] P. Grambsch, "Sequential sampling based on the observed Fisher information to guarantee the accuracy of the maximum likelihood estimator," *Ann. Statist.*, vol. 11, no. 1, pp. 68–77, 1983.
- [27] G. Fellouris, "Asymptotically optimal parameter estimation under communication constraints," *Ann. Statist.*, vol. 40, no. 4, pp. 2239–2265, Aug. 2012.
- [28] A. Wald, *Sequential Analysis*, Wiley, New York, NY, 1947.

- [29] H.V. Poor, *An Introduction to Signal Detection and Estimation*, Springer, New York, NY, 1994.